Strengthened Cutset Upper Bound on the Capacity of the Relay Channel and Applications

Abbas El Gamal
Dept. of Electrical Engineering
Stanford University
Email: abbas@ee.stanford.edu

Amin Gohari
Tehran Institute for Advanced Studies
Email: a.gohari@teias.institute

Chandra Nair
Dept. of Information Engineering
The Chinese University of Hong Kong
Email: chandra@ie.cuhk.edu.hk

2021 IEEE International Symposium on Information Theory (ISIT) | 978-1-5386-8209-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISIT45174.2021.9517975

***Abstract*—We establish a new upper bound on the capacity of the relay channel which is tighter than all previous bounds. The upper bound uses traditional weak converse techniques involving mutual information inequalities and identification of auxiliary random variables via past and future channel random variable sequences. We show that the new bound is strictly tighter than all previous bounds for the Gaussian relay channel for every set of non-zero channel gains. When specialized to the class of relay channels with orthogonal receiver components, the bound resolves a conjecture by Kim on a class of deterministic relay channels. When further specialized to the class of product-form relay channels with orthogonal receiver components, the bound resolves a generalized version of Cover's relay channel problem, recovers the recent upper bound for the Gaussian case by Wu et al. and also improves upon the recent bounds for the binary symmetric case by Wu et al. and Barnes et al., which were all obtained using non-traditional geometric proof techniques.**

A full version of this paper is accessible at: https://arxiv.org/pdf/2101.11139.pdf

I. INTRODUCTION

The relay channel, first introduced by van-der Meulen in 1971 [VDM71], is a canonical model of multi-hop communication networks in which a sender X wishes to communicate to a receiver Y with the help of a relay $(X_{\rm r}, Y_{\rm r})$ over a memoryless channel of the form $p(y, y_{\rm r}|x, x_{\rm r})$. The capacity of this channel, which is the highest achievable rate from X to Y, is not known in general. In [CEG79], lower bounds on the capacity, later termed decode-forward, partial decode-forward, and compress-forward, and the cutset upper bound were established. These bounds were shown to coincide for several special classes of channels, including degraded [CEG79], semi-deterministic [EGA82], and orthogonal sender components [EGZ05] relay channels. In [ARY09], the cutset bound was shown not to be tight in general via an example relay channel with orthogonal receiver components. These results and others are detailed in Chapter 19 of [EK11]. In a series of recent papers [WOX17], [WBO19], [LO19], motivated by Cover's problem concerning a relay channel with orthogonal receiver components [Cov87], new highly specialized upper bounds, which are also tighter than the cutset bound, are developed. While the bounds in [CEG79], [TU08], [ARY09] use standard weak converse techniques involving basic mutual information bounds and Gallager-type auxiliary random variable identification, the recent bounds in [WOX17], [WBO19], [LO19] for symmetric Gaussian and binary symmetric relay channels with orthogonal receiver components use more sophisticated arguments from convex geometry and functional analysis.

More recently, Gohari-Nair [GN20] developed a new upper bound on the capacity of the general relay channel and showed that it can be strictly tighter than the cutset bound. Their bound uses traditional converse techniques, including identification of auxiliary random variables using past and future channel variable sequences which has been used in several converse proofs, e.g., see [CK78], [EG79], and the new idea of auxiliary receiver. The upper bounds we present in this paper are natural extensions of the upper bound in [GN20]. Our bounds and their applications do not include an auxiliary receiver because we are not able to find an example with an auxiliary receiver that can strictly improve over the bound without it. We will focus our attention on the class of relay channels *without self-interference* $p(y_{\rm r}|x)p(y|x, x_{\rm r}, y_{\rm r})$ because they include and generalize several interesting relay channel settings that have been receiving significant attention in recent years.

Although the techniques used to establish the upper bounds in this paper have been employed in many previous works, the contributions of this paper are in the judicious manner in which these techniques are applied to obtain the tightest known upper bounds on the capacity of the relay channel and the rather nontrivial evaluations of these complex bounds to obtain tighter and more general bounds for several classes of relay channels.

In the following section we formally introduce the relay channel capacity problem and discuss our results (see Figure 1). While some of the proofs are given, the rest can be found in the full version of this paper at [EGGN21].

II. DEFINITIONS AND STATEMENT OF THE RESULTS

We adopt most of our notation from [EK11]. In particular, we use Y^i to denote the sequence $(Y_1, Y_2, \cdots, Y_i)$, and Y_i^j to denote $(Y_i, Y_{i+1}, \cdots, Y_j)$. Unless stated otherwise, logarithms are to the base 2. We use $p(x)$ to indicate the probability mass function of a discrete random variable X and P_Y to indicate the probability distribution of an arbitrary random variable Y.

The discrete memoryless relay channel consists of four alphabets $\mathcal{X}$, $\mathcal{X}_{\rm r}$, $\mathcal{Y}_{\rm r}$, $\mathcal{Y}$, and a collection of conditional pmfs $p(y_{\rm r}, y|x, x_{\rm r})$ on $\mathcal{Y}_{\rm r} \times \mathcal{Y}$. A $(2^{nR}, n)$ code for the discrete

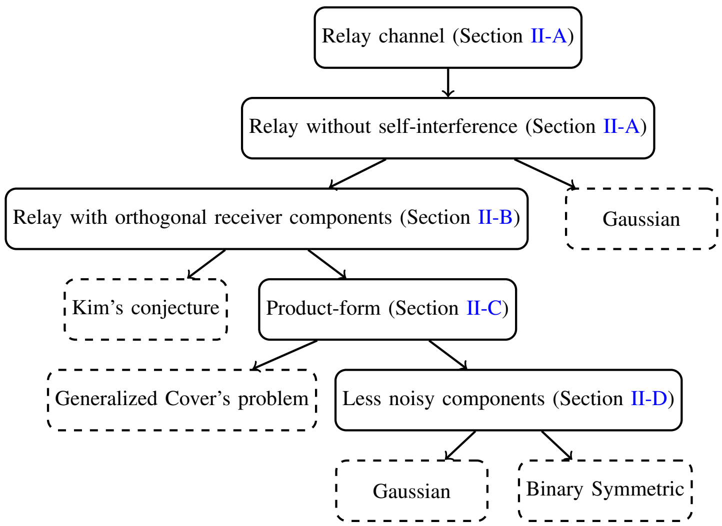

Fig. 1. Classes of relay channels for which upper bounds are established. An arrow from box A to box B indicates that class A includes class B. Dashed boxes indicate applications for which we compute the bounds.

memoryless relay channel $p(y, y_{\mathrm{r}}|x, x_{\mathrm{r}})$ consists of a message set $[1 : 2^{nR}]$, an encoder that assigns a codeword $x^n(m)$ to each message $m \in [1 : 2^{nR}]$, a relay encoder that assigns a symbol $x_{\mathrm{r}i}(y_{\mathrm{r}}^{i-1})$ to each past received sequence y_{r}^{i-1} for each time $i \in [1 : n]$, and a decoder that assigns an estimate $\hat{M}$ or an error message ε to each received sequence y^n. We assume that the message M is uniformly distributed over $[1 : 2^{nR}]$. The definitions of the average probability of error, achievability and capacity follow those in [EK11].

A. Upper bound for the relay channel without self-interference

We first consider the following class of relay channels.

Definition 1. A relay channel is said to be *without self-interference* if $p(y, y_{\mathrm{r}}|x, x_{\mathrm{r}}) = p(y_{\mathrm{r}}|x)p(y|x, x_{\mathrm{r}}, y_{\mathrm{r}})$.

We now present an upper bound for this class of relay channels. All the bounds discussed in the following sections follow from this bound either by relaxing the constraints or specializing it to subclasses of without self-interference relay channels.

Theorem 1. *Any achievable rate R for a general discrete memoryless relay channel $p(y_{\mathrm{r}}|x)p(y|x, x_{\mathrm{r}}, y_{\mathrm{r}})$ must satisfy the following inequalities*

$$\begin{aligned} R &\le I(X; Y, Y_{\mathrm{r}}|X_{\mathrm{r}}) - I(U; Y|X_{\mathrm{r}}, Y_{\mathrm{r}}), && (1)\\ R &\le I(X; Y, Y_{\mathrm{r}}|X_{\mathrm{r}}) - I(V; Y|X_{\mathrm{r}}, Y_{\mathrm{r}}) \\ &\qquad - I(X; Y_{\mathrm{r}}|V, X_{\mathrm{r}}, Y) && (2)\\ &= I(X; Y, V|X_{\mathrm{r}}) - I(V; X|X_{\mathrm{r}}, Y_{\mathrm{r}}), && (3)\\ R &\le I(X, X_{\mathrm{r}}; Y) - I(V; Y_{\mathrm{r}}|X_{\mathrm{r}}, X, Y), && (4)\end{aligned}$$

for some $p(u, x, x_{\mathrm{r}})p(y, y_{\mathrm{r}}|x, x_{\mathrm{r}})p(v|x, x_{\mathrm{r}}, y_{\mathrm{r}})$ satisfying

$$I(V, X_{\mathrm{r}}; Y_{\mathrm{r}}) - I(V, X_{\mathrm{r}}; Y) = I(U; Y_{\mathrm{r}}) - I(U; Y). \quad (5)$$

Further it suffices to consider $|\mathcal{V}| \le |\mathcal{X}||\mathcal{X}_{\mathrm{r}}||\mathcal{Y}_{\mathrm{r}}|+2$ and $|\mathcal{U}| \le |\mathcal{X}||\mathcal{X}_{\mathrm{r}}| + 1$.

The proof, given in [EGGN21], utilizes the following identification of the auxiliary random variables

$$V_i = (Y_{i+1}^n, Y_{\mathrm{r}}^{i-1}), \quad U_i = (Y_{\mathrm{r}i+1}^n, Y^{i-1}).$$

Remark 1 (Upper bound on the capacity of the general relay channel capacity.). An inspection of the proof of this theorem reveals that the without self-interference assumption is used only to establish the Markov chain $U \to X, X_{\mathrm{r}} \to Y, Y_{\mathrm{r}}$. If we relax this to $U \to X, X_{\mathrm{r}}, Y_{\mathrm{r}} \to Y$, then the same rate-constraints must hold for any relay channel.

Remark 2. It is immediate that the bound in Theorem 1 is at least as tight as the cut-set bound in [CEG79]

$$C \le \max_{p(x,x_{\mathrm{r}})} \{I(X_{\mathrm{r}}, X; Y), I(X; Y, Y_{\mathrm{r}}|X_{\mathrm{r}})\}. \quad (6)$$

Remark 3. From (3) and (4), we deduce that

$$\begin{aligned} R \le \max_{p(x,x_{\mathrm{r}})p(v|x,x_{\mathrm{r}},y_{\mathrm{r}})} \min\big\{&I(X; Y, V|X_{\mathrm{r}}) - I(V; X|X_{\mathrm{r}}, Y_{\mathrm{r}}),\\ &I(X, X_{\mathrm{r}}; Y) - I(V; Y_{\mathrm{r}}|X_{\mathrm{r}}, X, Y)\big\}.\end{aligned}$$

If we replace the maximum over $p(x, x_{\mathrm{r}})p(v|x, x_{\mathrm{r}}, y_{\mathrm{r}})$ with maximum over $p(x)p(x_{\mathrm{r}})p(v|x_{\mathrm{r}}, y_{\mathrm{r}})$, we obtain the equivalent form of the compress-forward lower bound without time-sharing random variable Q in [EK11].

Remark 4. If the cut-set bound is tight for a relay channel of the form $p(y, y_{\mathrm{r}}|x, x_{\mathrm{r}}) = p(y|x, x_{\mathrm{r}})p(y_{\mathrm{r}}|x)$ and capacity achieves the MAC bound in the cut-set bound, i.e., $C = I(X_{\mathrm{r}}, X; Y)$ for the maximizing $p(x, x_{\mathrm{r}})$, then capacity is achievable by partial-decode-forward, since from (4) we obtain that $I(V; Y_{\mathrm{r}}|X_{\mathrm{r}}, X, Y) = 0$. The assumption $p(y, y_{\mathrm{r}}|x, x_{\mathrm{r}}) = p(y|x, x_{\mathrm{r}})p(y_{\mathrm{r}}|x)$ implies that $I(V; Y_{\mathrm{r}}|X_{\mathrm{r}}, X) = 0$. The constraint $V \to X, X_{\mathrm{r}} \to Y_{\mathrm{r}}$, then implies that

$$\begin{aligned} &I(X; Y, Y_{\mathrm{r}}|X_{\mathrm{r}}) - I(V; Y|X_{\mathrm{r}}, Y_{\mathrm{r}}) - I(X; Y_{\mathrm{r}}|V, X_{\mathrm{r}}, Y)\\ &= I(V; Y_{\mathrm{r}}|X_{\mathrm{r}}) + I(X; Y|V, X_{\mathrm{r}}),\end{aligned}$$

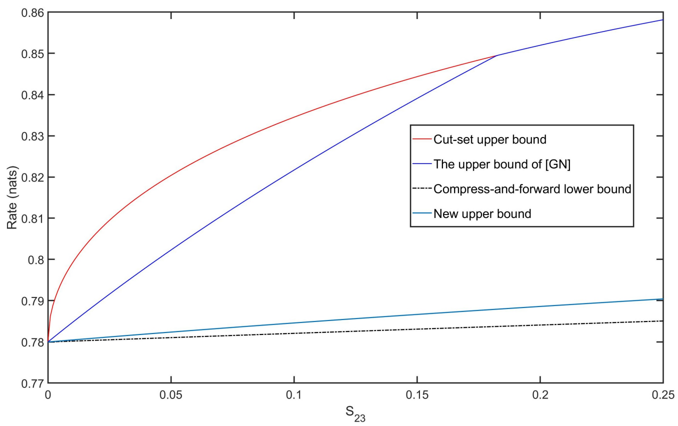

Fig. 2. Plots of the bounds for the Gaussian relay channel with $S_{13} = 3.7585$, $S_{12} = 1.2139$. The new upper bound is a weakened version of Theorem 1.

which is the partial-decode-forward lower bound. An application of this observation to solve a concrete question, which has the same flavor as the one introduced by Cover in [Cov87], can be found in [EGGN21].

Gaussian relay channel. The Gaussian relay channel is defined by

$$\begin{aligned} Y_{\mathrm{r}} &= g_{12}X + Z_{\mathrm{r}}, \\ Y &= g_{13}X + g_{23}X_{\mathrm{r}} + Z, \end{aligned} \tag{7}$$

where g_{12}, g_{13}, and g_{23} are channel gains, and $Z \sim \mathcal{N}(0,1)$ and $Z_{\mathrm{r}} \sim \mathcal{N}(0,1)$ are independent noise components. We assume average power constraint P on each of X and X_{r}.

Remark 5. Note that as defined, the Gaussian relay channel belongs to the class of relay channels without self-interference.

Let $S_{12} = g_{12}^2 P$, $S_{13} = g_{13}^2 P$ and $S_{23} = g_{23}^2 P$. In [EGGN21], a weakened version of the upper bound in Theorem 1, which is easier to evaluate for the Gaussian relay channel, is given. Figure 2 compares this weakened bound with the bound in [GN20, Proposition 1] for a scalar Gaussian relay channel.

Theorem 2. *The bound in Theorem 1 for the Gaussian relay channel is strictly tighter than the cut-set upper bound for every non-zero values of g_{12}, g_{13}, g_{23}.*

The idea of the proof is as follows, see [EGGN21] for details. Let C_{CS} denote the cut-set bound. Assume that C_{CS} is tight, that is,

$$\max_{P_{X,X_{\mathrm{r}}}:\, \mathrm{E}(X^2)\le P,\, \mathrm{E}(X_{\mathrm{r}}^2)\le P} \min\{I(X,X_{\mathrm{r}};Y), I(X;Y,Y_{\mathrm{r}}|X_{\mathrm{r}})\}$$

is achievable. We know (see Section 16.5 of [EK11]) that the maximum is attained via a unique jointly Gaussian distribution and $C_{\mathrm{CS}} = I(X;Y,Y_{\mathrm{r}}|X_{\mathrm{r}})$ holds. If $C = C_{\mathrm{CS}}$ from Theorem 1, we deduce the existence of a $P_{V|X,X_{\mathrm{r}},Y_{\mathrm{r}}}$ such that

$$I(V;Y|X_{\mathrm{r}},Y_{\mathrm{r}}) = I(X;Y_{\mathrm{r}}|V,X_{\mathrm{r}},Y) = 0.$$

We show that the above equality cannot occur.

B. Relay channels with orthogonal receiver components

In this section we present results for the following sub-class of relay channels without self-interference.

Definition 2. A relay channel is said to be *with orthogonal receiver components* (also referred to as *primitive*) (see Section 16.7.3 in [EK11]) if $Y = (Y_1, Y_2)$, where $p(y_1, y_2, y_{\mathrm{r}}|x, x_{\mathrm{r}}) = p(y_1, y_{\mathrm{r}}|x)p(y_2|x_{\mathrm{r}})$. It is known that the capacity of the above relay channel depends only on the capacity of the channel $p(y_2|x_{\mathrm{r}})$, hence we can substitute the relay-to-receiver channel $p(y_2|x_{\mathrm{r}})$ with a noiseless link of the same capacity C_0 [Kim07].

The following provides an equivalent characterization of the upper bound in Theorem 1 for relay channels with orthogonal receiver components.

Proposition 1. *Any achievable rate R for the relay channel with orthogonal receiver components $p(y_1, y_{\mathrm{r}}|x)$ with a relay-to-receiver link of capacity C_0 must satisfy the following inequalities*

$$R \le I(X;Y_1,Y_{\mathrm{r}}) - I(U;Y_1|Y_{\mathrm{r}}), \tag{8}$$

$$\begin{aligned} R &\le I(X;Y_1,Y_{\mathrm{r}}) - I(V;Y_1|Y_{\mathrm{r}}) - I(X;Y_{\mathrm{r}}|V,Y_1) \quad (9) \\ &= I(X;Y_1,V) - I(V;X|Y_{\mathrm{r}}), \quad (10) \end{aligned}$$

$$R \le I(X;Y_1) + C_0 - I(V;Y_{\mathrm{r}}|X,Y_1) \tag{11}$$

for some $p(u,x)p(y_1,y_{\mathrm{r}}|x)p(v|x,y_{\mathrm{r}})$ such that

$$\begin{aligned} I(U;Y_{\mathrm{r}}) - I(U;Y_1) &\le I(V;Y_{\mathrm{r}}) - I(V;Y_1) \\ &\le I(U;Y_{\mathrm{r}}) - I(U;Y_1) + C_0. \end{aligned}$$

Further it suffices to consider $|\mathcal{V}| \le |\mathcal{X}||\mathcal{Y}_{\mathrm{r}}| + 2$ and $|\mathcal{U}| \le |\mathcal{X}| + 1$.

Kim's conjecture. We use Proposition 1 to prove a conjecture posed by Kim in [Kim07, Question 2] for a class of of deterministic relay channels with orthogonal receiver components described by $p(y_1, y_{\mathrm{r}}|x)$, where $X = f(Y_1, Y_{\mathrm{r}})$ for some function f.

Theorem 3. *Let $\mathcal{C}(C_0)$ be the supremum of achievable rates R for a given C_0. Let C_0^* be the minimum value of C_0 for which $\mathcal{C}(C_0) = \mathcal{C}(\infty) = \log|\mathcal{X}|$. Then $C_0^* = H_G(Y_{\mathrm{r}}|Y_1)$ and is achieved by a uniform distribution on X. Here $H_G(Y_{\mathrm{r}}|Y_1)$ denotes the conditional graph entropy of the characteristic graph of (Y_{r}, Y_1) and the function f (as defined in [OR95]).*

The proof of this theorem is given in Section III-A.

C. Product-form relay channels

Consider the following class of relay channels with orthogonal receiver components.

Definition 3. A relay channel with orthogonal receiver components is said to be *product-form* if $p(y_1, y_{\mathrm{r}}|x) = p(y_1|x)p(y_{\mathrm{r}}|x)$.

Generalized Cover Relay Channel Problem. We will need the following definitions.

Definition 4. A discrete-memoryless channel $p(y|x)$ is said to be *generic* if the channel matrix, P with entries $P_{x,y} = p(y|x)$, $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ is full row rank.

Remark 6. It is immediate that if $p(y_1|x_1)$ and $p(y_2|x_2)$ are generic, then so is $p(y_1|x_1) \otimes p(y_2|x_2)$. That is, the class of generic channels is closed under a product operation.

Definition 5. A product-form relay channel is said to be *generic* if the channel $p(y_1|x)$ is generic.

In [Cov87], Cover posed a special (symmetric) case of the following problem: Consider a generic product-form relay channel and let $\mathcal{C}(C_0)$ be the supremum of achievable rates R for a given C_0. What is the critical value C_0^* for which $C_0^* = \inf\{C_0 : \mathcal{C}(C_0) = \mathcal{C}(\infty) = \max_{p(x)} I(X; Y_1, Y_\mathrm{r})\}$?

This problem has recently attracted a fair amount of attention and non-traditional methods have been used to answer the question as well as obtain new upper bounds for $\mathcal{C}(C_0)$ for Gaussian channels and binary-symmetric channels. As we will show in the next subsections our new upper bound, which uses traditional converse techniques, recovers and (in the binary-symmetric case) improves on these recent results. In this section, we show that we can answer the generalized Cover relay channel problem.

We can answer the generalized Cover's open problem by evaluating the bound in Proposition 1.

Theorem 4. *Let C_0^* be the minimum value of C_0 such that $\mathcal{C}(C_0) = \mathcal{C}(\infty) = \max_{p(x)} I(X; Y_1, Y_\mathrm{r})$ for a generic product-form relay channel and R_0^* be the minimum value of C_0 such that $R_{\mathsf{CF}}(C_0) = \mathcal{C}(\infty) = \max_{p(x)} I(X; Y_1, Y_\mathrm{r})$ for the same relay channel. Then $C_0^* = R_0^*$.*

The proof of this theorem is given in Section III-B.

Remark 7. During the finalization of this manuscript the authors became aware of [Liu20] which uses results and techniques in convex geometry to arrive at a solution for Theorem 4.

D. Product-form relay channels with less noisy components

Consider the following class of product-form relay channels.

Definition 6. A product-form relay channel described by $p(y_\mathrm{r}|x)p(y_1|x)$ and a relay-to-receiver link of capacity C_0 is said to have less-noisy components if it satisfies the conditions:

$$I(U; Y_\mathrm{r}) \leq I(U; Y_1) \text{ for every } p(u, x). \tag{12}$$

For this class of relay channels, we can specialize Proposition 1 to obtain the following bound.

Proposition 2. *Any achievable rate R for product-form relay channel with less-noisy components must satisfy the following conditions*

$$R \leq I(X; Y_1, Y_\mathrm{r}) - I(V; Y_1|Y_\mathrm{r}) - I(X; Y_\mathrm{r}|V, Y_1) \tag{13}$$

for some $p(x)p(y_1, y_\mathrm{r}|x)p(v|x, y_\mathrm{r})$ satisfying

$$I(V; Y_\mathrm{r}) - I(V; Y_1) \leq C_0. \tag{14}$$

Further it suffices to consider $|\mathcal{V}| \leq |\mathcal{X}||\mathcal{Y}_\mathrm{r}| + 1$.

In the following we will also refer to the following class of relay channels.

Definition 7. A product-form relay channel is said to be *symmetric* if $\mathcal{Y}_1 = \mathcal{Y}_\mathrm{r}$, and $p_{Y_1|X}(y|x) = p_{Y_\mathrm{r}|X}(y|x)$ for all x, y.

It immediately follows that if a product-form relay channel is symmetric, then it is less noisy, and Proposition 2 provides an upper bound for the symmetric class.

Gaussian Product-form relay channel with less-noisy components. Consider a Gaussian relay channel with orthogonal receiver components described by

$$Y_1 = X + W_1,$$
$$Y_\mathrm{r} = X + W_\mathrm{r},$$

where $W_1 \sim \mathcal{N}(0, N_1)$ and $W_\mathrm{r} \sim \mathcal{N}(0, N_\mathrm{r})$ are independent of each other and of X, and a link of capacity C_0 from the relay to the destination. We assume average power constraint P on X and define $S_{12} = P/N_\mathrm{r}$, $S_{13} = P/N_1$ and $S_{23} = 2^{2C_0} - 1$.

For $S_{13} \geq S_{12}$, the relay channel has less-noisy components and the upper bound in Proposition 2 reduces to the following.

Proposition 3. *Any achievable rate R for the Gaussian product-form relay channel with $S_{13} \geq S_{12}$ must satisfy the following*

$$R \leq \frac{1}{2} \log \left(1 + S_{13} + \frac{S_{12}(S_{13}+1)S_{23}}{(S_{13}+1)(S_{23}+1) - 1} \right). \tag{15}$$

Remark 8. Note that the compress-forward lower bound for this relay channel as given in [EK11, Eq. 16.17] implies that

$$C \geq \frac{1}{2} \log \left(1 + S_{13} + \frac{S_{12}(S_{13}+1)S_{23}}{S_{12} + (S_{13}+1)(S_{23}+1)} \right). \tag{16}$$

Furthermore, the above lower bound can be improved by time-sharing at the transmitter [EK11, Sec. 16.8] or at the relay [WBO19, Footnote 2].

Figure 3 depicts the upper bound in Proposition 3 along with the cut-set upper bound and the compress-forward lower bound.

Remark 9. In [WBO19], [WBÖ17], an upper bound on the capacity of the Gaussian product-form relay channel is established (see also [BWÖ20] for an alternative proof). Although the techniques used to prove this theorem are completely different from those used in this paper, it turns out quite surprisingly that bound (15) coincides with the bound in Proposition 3 for the symmetric special case. In simulations, our bound in Proposition 3 coincides with the bound in [WBÖ17] for $S_{13} \geq S_{12}$.

Symmetric binary relay channel with orthogonal receiver components. Assume that the relay channel with orthogonal receiver components is described by $p(y_1, y_\mathrm{r}|x)$ and a link of rate C_0 from relay to the destination such that $p(y_1, y_\mathrm{r}|x) = p(y_1|x)p(y_\mathrm{r}|x)$, where $x, y_1, y_\mathrm{r} \in \{0, 1\}$. Moreover, assume

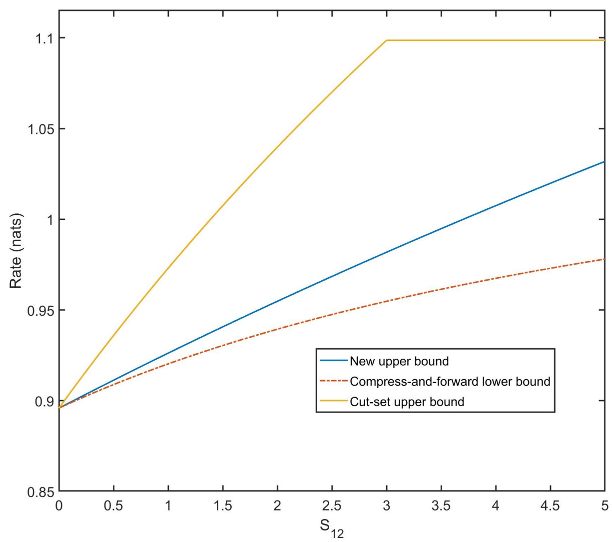

Fig. 3. Plots of the bounds for the Gaussian product-form relay channel with $S_{13} = 5$, $S_{23} = 0.5$.

that both the channels $p(y_1|x)$ and $p(y_r|x)$ are binary symmetric channels with crossover probability $\rho \in [0, 1/2]$.

By specializing Proposition 2 to this channel and using symmetry properties of the channel, we obtain the following bound.

Theorem 5. *Given arbitrary $\lambda \in [0,1]$ and $c \in [0,1]$, let $g_\lambda(c)$ be the maximum of $(1-\lambda)(H(Y_1) - H(Y_r)) + H(Y_r|X)$ over all joint probability distributions $p(x, y_r)$ on $\{0,1\} \times \{0,1\}$ satisfying $p(x, y_r)(0,1) + p(x, y_r)(1,0) = c$. For any fixed $\lambda \in [0,1]$, let $\mathscr{C}[g_\lambda] : [0,1] \mapsto \mathbb{R}$ be the upper concave envelope of the function $g_\lambda(\cdot)$, i.e., the smallest concave function that dominates $g_\lambda(\cdot)$ from above. Any achievable rate R for a symmetric binary relay channel with orthogonal receiver components with parameter ρ must satisfy the following*

$$R \leq 1 - 2H_2(\rho) + \lambda C_0 + \mathscr{C}[g](\rho)$$

for any $\lambda \in [0,1]$, where $H_2(x) = -x\log(x) - (1-x)\log(1-x)$ is the binary entropy function.

Figure 4 shows that our new upper bound strictly improves upon the bounds given in [WOX17] and [BWÖ17].

III. Proof of some of the results

A. Proof of Theorem 3

As argued in [Kim07], $C_0^* \leq H_G(Y_r|Y_1)$. It remains to show that $C_0^* \geq H_G(Y_r|Y_1)$. Let $C_0 = C_0^*$. From (9), we deduce that $I(V; Y_1|Y_r) = 0$. Since X is a function of (Y_1, Y_r), we deduce that $I(V; X|Y_r) = 0$. Consequently, (10) and (11) imply that the achievable rate $R = \log|\mathcal{X}|$ must satisfy the condition

$$R \leq \max_{p(x)p(v|y_r)} \min\{I(X; V, Y_1), \quad I(X; Y_1) + C_0 - I(V; Y_r|X, Y_1)\}.$$

This matches the rate achieved by compress-forward (see [EK11, Eq. 16.14]). Thus, compress-forward must achieve the

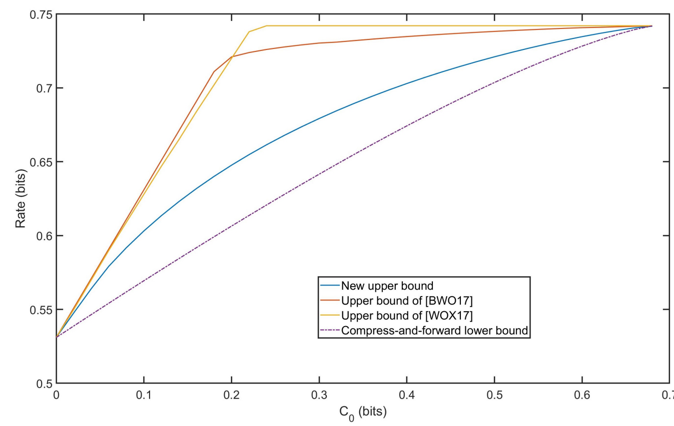

Fig. 4. Plots of the minimum of the two upper bounds give in [WOX17], the upper bound given in [BWÖ17], our new bound and the compress-forward lower bound for a symmetric binary relay channel with orthogonal receiver components with parameter $\rho = 0.1$.

rate $\log|\mathcal{X}|$ when $C_0 = C_0^*$. Using the characterization of the compress-forward lower bound in Proposition 3 of [Kim07], we obtain that $\log|\mathcal{X}| = I(X; \tilde{V}, Y_1)$ for some $p(\tilde{v}|y_r)$ such that $I(\tilde{V}; Y_r|Y_1) \leq C_0^*$. Therefore, X is uniformly distributed and $H(X|\tilde{V}, Y_1) = 0$. From [OR95, Theorem 2], we deduce that $C_0^* \geq H_G(Y_r|Y_1)$. This confirms Kim's conjecture in [Kim07].

B. Proof of Theorem 4

The result follows immediately from the definition that $R_0^* \geq C_0^*$. Therefore it suffices to show that $R_0^* \leq C_0^*$. Let C_0 be such that $\mathcal{C}(C_0) = \mathcal{C}(\infty) = \max_{p(x)} I(X; Y_1, Y_r)$. From the constraint (9) of the upper bound in Proposition 1, it follows that if $\mathcal{C}(\infty) = \max_{p(x)} I(X; Y_1, Y_r)$ is achievable, for a maximizing distribution $p^*(x)$, there exists a distribution $p(v|x, y_r)$ such that $I(V; Y_1|Y_r) = 0$. Since the channel $p_{Y_1|X}$ is generic, we show in [EGGN21] that $I(V; Y_1|Y_r) = 0$ implies that $I(V; X|Y_r) = 0$. Therefore $V \to Y_r \to X \to Y_1$ form a Markov chain. Then, the constraints in (10) and (11) imply that the rate R is achievable by compress-forward with the compression random variable V. Here, we utilize the characterization of the compress-forward for the relay channel with orthogonal receiver components given in [EK11, Eq. 16.14]. Consequently, we have that $R_{\mathsf{CF}}(C_0) = \mathcal{C}(C_0) = \mathcal{C}(\infty)$. Since this holds for any C_0 such that $\mathcal{C}(C_0) = \mathcal{C}(\infty) = \max_{p(x)} I(X; Y_1, Y_r)$, we have that $R_0^* \leq C_0^*$. This completes the proof.

Conclusion

We established a new outer bound for the relay channel and used it to obtain several new results, including showing that the cutset bound for the scalar Gaussian relay is suboptimal for all non-zero values of the parameters, providing a resolution of Kim's conjecture and a generalization of Cover's open problem, strictly improving upon previous outer bounds on the capacity of the symmetric binary relay channels with orthogonal receiver components.

REFERENCES

[ARY09] Marko Aleksic, Peyman Razaghi, and Wei Yu, *Capacity of a class of modulo-sum relay channels*, IEEE Transactions on Information Theory **55** (2009), no. 3, 921–930.

[BWÖ17] Leighton Pate Barnes, Xiugang Wu, and Ayfer Özgür, *A solution to cover's problem for the binary symmetric relay channel: Geometry of sets on the hamming sphere*, 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE, 2017, pp. 844–851.

[BWÖ20] Yikun Bai, Xiugang Wu, and Ayfer Özgür, *Information constrained optimal transport: From talagrand, to marton, to cover*, 2020 IEEE International Symposium on Information Theory (ISIT), IEEE, 2020, pp. 2210–2215.

[CEG79] Thomas Cover and Abbas El Gamal, *Capacity theorems for the relay channel*, IEEE Transactions on information theory **25** (1979), no. 5, 572–584.

[CK78] Imre Csiszár and Janos Korner, *Broadcast channels with confidential messages*, IEEE transactions on information theory **24** (1978), no. 3, 339–348.

[Cov87] Thomas Cover, *The capacity of the relay channel*, Open Problems in Communication and Computation, Springer, 1987, pp. 72–73.

[EG79] Abbas El Gamal, *The capacity of a class of broadcast channels*, IEEE Transactions on Information Theory **25** (1979), no. 2, 166–169.

[EGA82] Abbas El Gamal and Mohammad-Reza Aref, *The capacity of the semideterministic relay channel (corresp.)*, IEEE Transactions on Information Theory **28** (1982), no. 3, 536–536.

[EGGN21] Abbas El Gamal, Amin Gohari, and Chandra Nair, *Strengthened cutset upper bounds on the capacity of the relay channel and applications*, arXiv preprint arxiv:2101.11139 (2021).

[EGZ05] Abbas El Gamal and Sina Zahedi, *Capacity of a class of relay channels with orthogonal components*, IEEE Transactions on Information Theory **51** (2005), no. 5, 1815–1817.

[EK11] Abbas El Gamal and Young-Han Kim, *Network information theory*, Cambridge University Press, 2011.

[GN20] Amin Gohari and Chandra Nair, *Outer bounds for multiuser settings: The auxiliary receiver approach*, submitted to the IEEE Transactions on Information Theory (2020).

[Kim07] Young-Han Kim, *Coding techniques for primitive relay channels*, Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput, 2007, p. 2007.

[Liu20] Jingbo Liu, *Minorization via mixed volumes and cover's problem for general channels*, arXiv preprint arXiv:2012.14521 (2020).

[LO19] Jingbo Liu and Ayfer Ozgur, *New converses for the relay channel via reverse hypercontractivity*, 2019 IEEE International Symposium on Information Theory (ISIT), July 2019, pp. 2878–2882.

[OR95] Alon Orlitsky and James R Roche, *Coding for computing*, Proceedings of IEEE 36th Annual Foundations of Computer Science, IEEE, 1995, pp. 502–511.

[TU08] Ravi Tandon and Sennur Ulukus, *A new upper bound on the capacity of a class of primitive relay channels*, 2008 46th Annual Allerton Conference on Communication, Control, and Computing, IEEE, 2008, pp. 1562–1569.

[VDM71] Edward C Van Der Meulen, *Three-terminal communication channels*, Advances in applied Probability (1971), 120–154.

[WBÖ17] Xiugang Wu, Leighton Pate Barnes, and Ayfer Özgür, *The geometry of the relay channel*, 2017 IEEE International Symposium on Information Theory (ISIT), IEEE, 2017, pp. 2233–2237.

[WBO19] Xiugang Wu, Leighton Pate Barnes, and Ayfer Ozgur, *"the capacity of the relay channel": Solution to cover's problem in the gaussian case*, IEEE Transactions on Information Theory **65** (2019), no. 1, 255–275.

[WOX17] Xiugang Wu, Ayfer Ozgur, and Liang-Liang Xie, *Improving on the cut-set bound via geometric analysis of typical sets*, IEEE Transactions on Information Theory **63** (2017), no. 4, 2254–2277.